Funktionalableitungen „in a nutshell“

Michael J. Gruber

1 Endlichdimensionale Analogie

Ist $L : V \to \mathbb{R}$ eine (nichtlineare) Abbildung auf einem endlichdimensionalen Vektorraum, so ist die Ableitung von L in $v \in V$ in Richtung $h \in V$ gegeben durch

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left(L(v + \varepsilon h) - L(v) \right) =: \frac{dL(v)}{dv}(h)$$

Die Existenz des Limes ist gleichbedeutend mit der Existenz der Ableitung. Diese ist linear in h, man kann sie also durch ein Skalarprodukt ausdrücken, und schreibt einfach

$$\frac{dL(v)}{dv}(h) = \frac{dL(v)}{dv} \cdot h$$

mit dem Skalarprodukt „$\cdot$“. In diesem Sinne ist

$$\frac{dL(v)}{dv} = \nabla L = \mathrm{grad} L$$

2 Lineare Funktionale

Ist V ein unendlichdimensionaler Vektorraum mit einer Topologie (damit man weiß, was „stetig“ heißt), so ist nicht unbedingt jede stetige lineare Abbildung $V \to \mathbb{R}$ mit einem Vektor aus V identifizierbar. Dies geht nur, wenn V ein Hilbert-Raum ist.

Beispiel: Ist $V = C_c^\infty(\mathbb{R}^d)$ (glatte Funktionen mit kompaktem Träger auf $\mathbb{R}^d$) und ist $g \in V$, so ist die Abbildung

$$T_g : f \mapsto \int f(x) g(x)\, dx$$

linear und stetig. In diesem Sinne kann man jede Funktion g auch als lineares Funktional T_g auffassen.

Beispiel: Ist $V = C_c^\infty(\mathbb{R}^d)$, so ist die Abbildung

$$\delta_{x_0} : f \mapsto f(x_0)$$

linear und stetig (Auswertung im Punkt x_0). Die Anwendung des linearen Funktionals δ_{x_0} auf f bezeichnet man mit $\delta_{x_0}[f]$ (δ-Distribution oder Dirac-Distribution). In Analogie zum Endlichdimensionalen schreibt man

$$\delta_{x_0}[f] = \langle \delta_{x_0}, f \rangle = \int \delta_{x_0} f = \int \delta_{x_0}(x) f(x)\, dx = \int \delta(x - x_0) f(x)\, dx,$$

da das Integral ein Skalarprodukt auf dem Funktionenraum ist. Man überzeugt sich aber leicht, daß $\delta(x - x_0)$ keine Funktion aus V sein kann, d.h. es gibt kein $g \in V$ so daß $T_g = \delta_{x_0}$.
Lineare stetige Funktionale auf Funktionenräumen nennt man auch Distributionen. Für weitere Beispiele und Details siehe Funktionalanalysis.

3 Distributive Ableitung

Ist V ein Raum differenzierbarer Funktionen (C^1 oder höher, kompakte Träger oder Schwarzraum, ...) auf $\mathbb{R}$ und T_g die Distribution wie oben, so kann man partiell integrieren:

$$T_{g'}[f] = \int g'(x)f(x)\,dx = -\int g(x)f'(x)\,dx = -T_g[f']$$

Daher definiert man für jede Distribution T durch

$$T'[f] := -T[f']$$

die distributive Ableitung von T, so daß insbesondere $(T_g)' = T_{g'}$. Man schreibt auch $T' = \partial_x T$.
Jede Distribution ist so differenzierbar. Insbesondere ist jede Funktion (distributiv) differenzierbar! Allerdings muß die Ableitung keine Funktion mehr sein.
Beispiel: Ist $\theta(x)$ die Heaviside-Funktion, so ist

$$\theta'[f] = -\theta[f'] = -\int \theta(x)f'(x)\,dx = -\int_0^\infty f'(x)\,dx = f(0) = \delta_0[f]$$

Die Ableitung der Heaviside-Funktion ist also die δ-Distribution.
Beispiel: Die Ableitung von δ_{x_0} berechnet man wie folgt aus der Definition:

$$\delta'_{x_0}[f] = -\delta_{x_0}[f'] = -f'(x_0)$$

Induktiv definiert man die höheren Ableitungen; analog distributive partielle Ableitungen für Distributionen auf $\mathbb{R}^d$ statt $\mathbb{R}$.

4 Funktionalableitung

Funktionale sind nichtlineare Abbildungen von einem Funktionenraum V nach $\mathbb{R}$. Eine sinnvolle Ableitung sollte daher eine Linearisierung in einem „Punkt“ (d.h. in einer Funktion f) sein, die ein Funktional ist — also ein lineares Funktional. Ist $L : V \to \mathbb{R}$ ein Funktional, so definiert man daher ganz wie beim Gradienten die Ableitung von L in $f \in V$ in Richtung $h \in V$ durch

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left(L[f + \varepsilon h] - L[f] \right) =: \frac{\delta L[f]}{\delta f}[h]$$

Die Existenz (und Stetigkeit) des Limes ist gleichbedeutend mit der Existenz (und Stetigkeit) der Ableitung. Diese ist linear in h (ein lineares Funktional), aber man kann sie nicht unbedingt durch ein Skalarprodukt mit einem Vektor aus V ausdrücken (siehe Abschnitt 2). Wegen der Linearität schreibt man einfach $\frac{\delta L[f]}{\delta f}$ für das Funktional (wie beim Gradienten).
Mehr Details unter dem Stichwort „Fréchetableitung“ in der Funktionalanalysis-Literatur. Man nennt die Funktionalableitung auch Variaton von $L[f]$ nach f oder Variationsableitung.

5 Beispiele und Rechenregeln

5.1 Lineares Funktional

Ist L selbst linear, so ist $L[f+\varepsilon h]-L[f]=\varepsilon L[h]$. Also ist $\frac{\delta L[f]}{\delta f}[h]=L[h]$ oder auch $\frac{\delta L[f]}{\delta f}=L$, wie man es von einer linearen Abbildung erwartet. Insbesondere: Wenn $L[f]=\int fg=T_g[f]$ ist mit einem festen g, so ist $\frac{\delta L[f]}{\delta f}[h]=L[h]=\int hg$. Man schreibt daher

$$\frac{\delta L[f]}{\delta f}=\frac{\delta T_g[f]}{\delta f}=T_g=g$$

im gleichen Sinne wie bei der δ-Distribution (Identifikation von g und T_g).

5.2 Kettenregel global

Sei $L[f]=\int F(f(x))\,dx$ für eine Funktion $F:\mathbb{R}\to\mathbb{R}$. Ist F einmal stetig differenzierbar, so ist:

$$\begin{aligned}\frac{1}{\varepsilon}(L[f+\varepsilon h]-L[f]) &= \frac{1}{\varepsilon}\int (F(f(x)+\varepsilon h(x))-F(f(x)))\,dx\\ &= \frac{1}{\varepsilon}\int \big(F'(f(x))\varepsilon h(x)+\mathcal{O}(\varepsilon^2 h(x)^2)\big)\,dx\\ &\to \int F'(f(x))h(x)\,dx\end{aligned}$$

Wir haben also $\dfrac{\delta L[f]}{\delta f}=F'(f)$.

Anwendung: In der Kontinuumsmechanik benutzt man $\dfrac{\delta}{\delta y}\displaystyle\int \tfrac{1}{2}(y(x))^2\,dx=y$ für quadratische potentielle Energie und $\dfrac{\delta}{\delta \dot y}\displaystyle\int \tfrac{1}{2}(\dot y(x))^2\,dx=\dot y$ für die kinetische Energie. $\dot y$ ist hier eine Funktion, die zunächst unabhängig von y ist, genauso wie in der Lagrange-Punktmechanik x und $\dot x$ voneinander unabhängige Koordinaten im Phasenraum sind.

5.3 Delta

Ist $L[f]=\delta_{x_0}[f]=f(x_0)$, so ist L linear! Nach 5.1 ist also

$$\frac{\delta\delta_{x_0}[f]}{\delta f}=\delta_{x_0},$$

was man in der Physik auch gerne so schreibt:

$$\frac{\delta f(x_0)}{\delta f(x)}=\delta(x-x_0)$$

Anschaulich gesprochen sind $f(x)$ und $f(x_0)$ unabhängig voneinander solange $x\neq x_0$.
Achtung: Nicht mit der distributiven Ableitung von δ verwechseln (siehe 3).
Anmerkung: Die Notation mit den gleichen δs für Variationsableitung und δ-Distribution ist leider so üblich.

5.4 Kettenregel lokal

Sei $L[f] = \delta_{x_0}[F(f)] = F(f(x_0))$ für eine Funktion $F : \mathbb{R} \to \mathbb{R}$. Wie bei der globalen Kettenregel (5.2) rechnet man nach:

$$\frac{\delta L[f]}{\delta f} = \frac{\delta\left(\delta_{x_0}[F(f)]\right)}{\delta f} = F'(f)\delta_{x_0}$$

Andere Notation:

$$\frac{\delta F(f(x_0))}{\delta f(x)} = \frac{\partial F(f)}{\partial f}\frac{\delta f(x_0)}{\delta f(x)} = F'(f(x))\delta(x - x_0)$$

Hier wird die Ähnlichkeit zur Kettenregel besonders deutlich. 5.2 ist gewissermaßen die integrierte Form. Durch Integrieren gelangt man von den lokalen zu den globalen Regeln.

5.5 Vertauschen mit ∂_x

Ist L ein Funktional, so definiert $K[f] := L[f']$ ein neues Funktional K, das nur noch von der Ableitung $f' =: \partial_x f$ abhängt. Die Funktionalableitung von K berechnet man wie folgt:

$$\begin{aligned}\frac{1}{\varepsilon}(K[f + \varepsilon h] - K[f]) &= \frac{1}{\varepsilon}(L[f' + \varepsilon h'] - L[f']) \\ &\to \frac{\delta L[f']}{\delta f'}[h'] = -\partial_x \frac{\delta L[f']}{\delta f'}[h]\end{aligned}$$

Im letzten Schritt haben wir die Definition der distributiven Ableitung aus Abschnitt 3 auf das lineare Funktional $\frac{\delta L[f']}{\delta f'}$ angewandt. Wir haben also:

$$\frac{\delta K[f]}{\delta f} = \frac{\delta L[f']}{\delta f} = -\partial_x \frac{\delta L[f']}{\delta f'}$$

Definiert man in Analogie zur distributiven Ableitung $L'[f] := -L[f']$, so ist

$$\frac{\delta L'[f]}{\delta f} = -\frac{\delta L[f']}{\delta f} = \partial_x \frac{\delta L[f']}{\delta f'},$$

d.h. distributive und Funktionalableitung vertauschen. Als Merkregel benutzt man gerne folgendes:

$$\frac{\delta f'(x_0)}{\delta f(x)} = \frac{\delta \frac{\partial f(x_0)}{\partial x_0}}{\delta f(x)} = \frac{\partial}{\partial x_0}\frac{\delta f(x_0)}{\delta f(x)} = \frac{\partial}{\partial x_0}\delta(x - x_0) = \delta'(x - x_0)$$

Anders gesagt: f und f' hängen voneinander ab. Das ist auch anschaulich klar, da die Funktion f die Ableitung f' festlegt.

Beispiel:

$$\frac{\delta}{\delta f}\int F(f'(x))\,dx \overset{5.5}{=} -\partial_x \frac{\delta}{\delta f'}\int F(f'(x))\,dx \overset{5.2}{=} -\partial_x F'(f') \overset{3}{=} -F''(f')f''$$

Anwendung: $\frac{\delta}{\delta y}\int \frac{1}{2}(y'(x))^2\,dx = -y''$ für Wechselwirkungsterme, wie sie z.B. bei der linearen Kette auftreten. Beachte, daß y und y' voneinander abhängen, nicht aber y und $\dot{y}$ (siehe 5.2).